Conference 2006

Rising Unemployment in the South African Labour Market: A Dynamic Analysis Using Birth Cohort Panels

Rulof Burger and Dieter von Fintel

Accelerated and Shared Growth in South Africa: Determinants, Constraints and Opportunities

18 - 20 October 2006

**The Birchwood Hotel and Conference Centre
Johannesburg, South Africa**

Conference organised with support from the EU

Rising Unemployment in the South African Labour Market:

A Dynamic Analysis Using Birth Cohort Panels

By Rulof Burger and Dieter von Fintel

Department of Economics, University of Stellenbosch

ABSTRACT

This paper attempts to take advantage of the wealth of cross-sectional household surveys conducted after South Africa's political transition, in order to gain some understanding of the dynamic nature of rising unemployment. A synthetic panel data approach is used to investigate, amongst other things, the source of the recent surge in youth unemployment. This is done by means of a decomposition analysis that identifies the between cohort and population group changes in the unemployment rate attributable to cyclical, generational and life-cycle effects. This analysis was also extended to a decomposition of the participation and employment rates. Our results indicate that the higher unemployment rates faced by the young are predominantly due to the disadvantage of entering the labour market more recently, rather than being attributable to their age. It was also shown that the bulk of this generational disadvantage was the result of an increase in the participation rate, rather than a decrease in employment opportunities. We find some correspondence between the cyclical variation in unemployment and the business cycle, although this relationship might be marred by survey-specific sampling error.

Employability in the South African Labour Market: Dynamic Evidence from Birth Cohort Panels

By Rulof Burger and Dieter von Fintel

Department of Economics, University of Stellenbosch

1 Introduction

Since the political transition in 1994, South Africa has experienced a large increase in its already high unemployment rate. Rising unemployment is a source of considerable concern to labour market participants and policymakers alike, and the benefits of better understanding the dynamic forces at play are potentially large. Since the inception of the October Household Surveys in 1994 and their successors, the Labour Force Surveys, the scope for labour market research has been extended substantially. This allows a clearer picture of the state of the labour market to emerge, but attempts to move beyond simple comparative static analysis has been plagued by difficulties.

In this paper we follow a synthetic panel approach in order to take advantage of the wealth of information available from all seventeen successive cross-sectional household surveys now available. It has been shown that South African unemployment has a strong age dimension. By following the mean characteristics of groups of individuals born in the same year from a pre-specified sub-population, the cohort panel methodology is ideally suited to tease out more information regarding the causes of this aspect of unemployment.

The paper is structured as follows. Section 2 outlines methodological and data issues. These are determined largely by the type of data at our disposal, namely successive cross sections. Due in part to concerns of data quality, this paper restricts itself to studying changes in the formal economy. Section 3 proceeds by decomposing unemployment into cyclical, generational and life-cycle components. The paper then investigates the respective roles that changes in the participation and employment rates play in driving each of unemployment's composite factors. Racial differentials are then investigated, with controls introduced in an attempt to explain the variation in cohort-specific fixed effects. Section 4 concludes the study.

2 Data and methodology

2.1 Methodology

Since the first October Household Surveys (OHS) in 1994, and the introduction of the more consistent Labour Force Surveys (LFS) in 2000, there has been a proliferation in studies that analyses South African unemployment. Differences in questionnaire design and sampling methodology, as well as the inconsistent derivation of labour market measures across surveys complicate direct comparisons of these surveys. However, taking a longer-term perspective mitigates some of these concerns. Much of the literature has therefore settled for comparative static analysis, by contrasting circumstances depicted by two cross-sections (for example Oosthuizen & Bhorat 2005, Casale and Posel 2005, Kingdon and Knight 2005). Coding problems make it difficult to identify the same households in different surveys, which precludes exploitation of the rotating panel design of the Labour Force Surveys.

It has been shown that South African unemployment has a strong age dimension, and youth unemployment has emerged as one of the most challenging social issues in the South African economy (see Mlatsheni & Rospabe 2002, amongst others). By following Deaton (1985) in constructing a birth cohort panel, we are able to study this feature of the increase in unemployment from a more dynamic perspective. By following the mean characteristics of groups of individuals born in the same year from a pre-specified sub-population (they may or may not be the same individuals in different surveys), it is possible to trace life cycle effects, in addition to the effect of business cycle fluctuations and longer-term structural trends.

The isolation of these effects allows a more focussed approach in consequent analyses. Is the increase in unemployment driven by a changing age-unemployment profile, or is it rather a problem that will afflict younger generations through-out their working lives? The answer to this question entails different approaches in tackling unemployment.

Following groups instead of individuals has both merits and drawbacks. The pseudo-panel approach introduces dynamic views, which are not otherwise possible. Variables are aggregated by the chosen cohorts. For instance, a participation dummy is averaged over all the individuals in a cohort (taking probability weights into account), and consequently represents the estimated participation rate for the specific cohort in each year under consideration. This in itself forms the basis for instructive descriptive analyses to trace, for example, the differences in unemployment paths across racial and gender boundaries (for one application of this method to South African labour market data, see Branson, 2005).

The primary benefit of this methodology is that it is unnecessary to follow the same individuals over time, but rather analyses the dynamics of "look alikes" (to use Deaton's (1985: 110) terminology). Important labour market questions, such as how the increase in unemployment affected blacks and whites of different ages can be answered without requiring information on the same individuals across time. As with panel data, it is possible to control for much of the unobserved heterogeneity that plague typical labour market studies. Unlike in a panel dataset, however, attrition is of little concern, since a set of individuals who meet the grouping criteria appear in each survey, despite the effects of migration, non-response and dissolution of households. A panel of semi-aggregated data allows for a richer analysis than either cross-sectional comparisons or pure time series data can offer: it is possible to lend a dynamic perspective to the investigation, yet maintain a breakdown of the composition of the variables under consideration. Deaton (1997: 117) applauds cohort data for providing a meeting point between disaggregated microeconomic information and variables' macroeconomic movements. These properties are exploited in this paper, by "disentangling the generational from life-cycle components" (Deaton 1997: 117), which is not possible in either the cross section or time series domains.

Unfortunately, this methodology does not provide an instrument to study individual transitions from one state to another (such as moving from the discouraged worker status to being an active searcher or finding employment), and is therefore not as informative as a pure panel dataset. Given the absense of properly constructed, nationally representative individual panels in South Africa, it is not clear that there is any way to address such issues. Furthermore, the sample average of a variable will not always be a good estimator of that cohort's population mean in each period. Should the mean have a large standard error or be constructed by only a few observations, it is questionable to invoke the law of large numbers, so that the sample averages could be plagued by considerable noise. This means that errors in variables are likely to occur, and that empirical models could suffer from bias and inconsistency.

If the data generating process is accurately characterised by the fixed effects model at an individual level, a cohort panel is unlikely to share the same time-invariant cohort-specific effect (Baltagi, 2005: 193). For a typical individual, the following may hold:

$$y_{it} = x_{it}^{'}\beta + \mu_i + v_{it} \quad t = 1, \dots, T$$

However, once aggregated by cohort, it is necessary to take cogniscance of the fact that different individuals constitute the same cohorts in different periods. The average of the individual fixed effects may therefore not be time-invariant. Hence, the "cohort version" of the fixed effects model, would, in its most general form be represented by:

$$\bar{y}_{ct} = \bar{x}_{it}^{'}\beta + \bar{\mu}_{ct} + \bar{v}_{ct} \quad c = 1, \dots, C; \;\; t = 1, \dots, T$$

where C is the number of cohorts, and T the number of time periods.

This model is only identified if it is assumed that $\bar{\mu}_{ct} = \bar{\mu}_c$, which demands that cohorts be constructed with a satisfactory number of individuals in each group. The same holds true for lagged dependent variables, since $y_{i\,t}$ is often determined by $y_{i\,t-1}$; clearly if the same individuals are not used to construct the means, measurement error could potentially conceal the autoregressive relationship (Hsiao, 2003: 284). Since cohort sample averages will converge on the population means for a large number of observations per cohort, we face a tradeoff between the benefit of additional degrees of freedom (by increasing the *number* of cohorts) against mitigting the errors-in-variables problem by choosing larger cohorts.

Deaton's (1985) initial work already proposed an adjusted fixed effects estimator, which scales each cohort by the square root of its constituent size and adjusts for the covariance structure of the means. This overcomes the measurement error bias which would usually result from applying a normal fixed effects estimator to synthetic panels. Unfortunately, this procedure complicates estimation, and may not result in considerable consistency gains. In practise, applied researchers often choose to ignore measurement error issues: it has been shown that with 100 or more observations within each cohort, bias is minimal and adjustments can be safely ignored (Verbeek and Nijman, 1992). In order for the standard fixed effects estimators to be valid, it is therefore important to choose cohorts with a sufficient number of observations. We start our empirical analysis by constructing only a birth cohort panel in section 3.1. Section 3.2 aggregates the data to a higher level, by also grouping according to race, which entails a reduction in cohort sizes. This raises the additional point that group sizes may not be identical or even similar across cohorts or time. The analysis of Indians, for example, becomes hazardous due to their small sample size. Inoue (2005) addresses this, along with efficiency and inferential concerns, by way of a GMM estimator. Since fixed effects estimators remain consistent with sufficient cohort sizes $(as\ C \rightarrow \infty)$ and degrees of freedom $(as\ N \rightarrow \infty)$, these issues are not considered in our analysis.

One of the benefits of cohort data, as mentioned above, is its ability to discern between life-cycle, generational and cyclical macroeconomic components of the dependent variable of interest. Deaton (1997: 123-127) outlines how a simple least square dummy variable regression (which is equivalent to the fixed effects estimator, though this method allows the cohort effects – the fixed effects in this scenario – to be directly estimated) can feasibly decompose unemployment, employment and participation into these respective elements. Time dummies are included to capture macroeconomic shocks, and their coefficient sizes can be compared to the business cycle to assess how responsive the labour market is (in terms of creating and shedding jobs) to fluctuations. Dummies for each birth year show how the "fixed effects" of each cohort behave: this gives us an indication of how the circumstances of various generations have changed. These cohort effects include, inter alia, the impact of *long-term* macroeconomic trends and changes in the average set of productive characteristics: simply said, these coefficients reveal the "employability" of different generations of South Africans, independent of the usual life-cycle effects. Controls are subsequently included, to absorb the explanatory power attributable to observable productive and demographic characteristics. The remaining profile offers a description of the unobserved cohort fixed effects, which measures, among other aspects, differences in educational quality, as well as the effects of the country's macroeconomic performance. The third set of dummies (representing age), isolates life-cycle effects. These separate out the "stylised facts" of the South African labour market, which functions independently of generational effects and the changing economic milieu. The usefulness of this technique is to account for possible sources of different types of unemployment. For instance, does youth unemployement arise because young people from every generation have always suffered this consequence in South Africa (life-cycle effect)? Or is youth unemployment the product of increasingly rigid labour markets or a declining quality of education, which erodes the skills base (cohort effects)? In answering these questions, decompositions are executed for the unemployment rate. The unemployment rate, u, can be expressed as

$$u \quad \frac{U}{L} \quad \frac{L \quad E}{L} \quad 1 \quad \frac{E}{L} \quad 1 \quad \frac{E}{P}\frac{P}{L} \quad 1 \quad \frac{E/P}{L/P} \quad 1 \quad \frac{e}{p}$$

where U is the number of unemployed individuals,

E is the number of employed individuals,

L is the labour force, and

P is the population of working age.

The unemployment rate is thus equal to 1 minus the ratio of the employment rate (e) to the participation rate (p), so that an increase in the unemployment rate can be the result of a decrease in the employment rate, an increase labour force participation, or a combination of the two. Therefore, the decomposition of unemployment is followed by a decomposition of its composite parts, the participation and employment rates. This effectively breaks down the above-mentioned contributions to unemployment further into participation changes (which could be accounted for by, for instance, policies regarding over-aged learners in the schooling system) coupled with the current absorptive capacity of the economy.

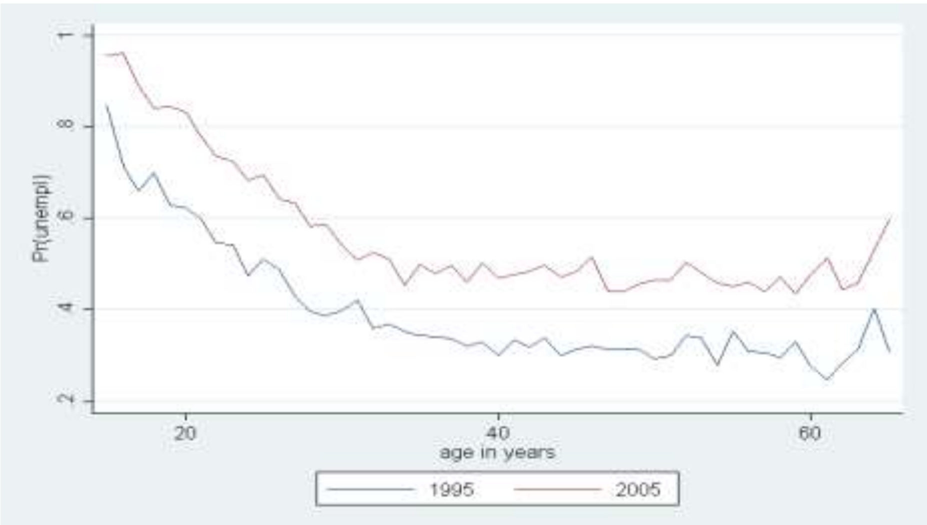
The practicalities of implementing this decomposition is an identification problem. All year, cohort and age effects should account largely for the dependent variable concerned. Perfect multicollinearity is a problem by definition, since age is a linear function of the current year and the birth year associated with a cohort. It is, however, possible to perform a simple transformation on the year dummies to estimate the equation subject to a zero restriction on the time effects (Deaton, 1997: 126). This makes intuitive sense, since these short-run macroeconomic fluctuations are assumed to average to zero in the long-run. The first age and cohort dummies are omitted to form bases, while a set of T - 2 new time dummies (omitting the first two years) are created by the following transformation:

$$y_t^* = y_t - \left[(t-1)y_2 - (t-2)y_1\right] \qquad t = 3, \dots, T$$

The time effects for the first and second years can subsequently be recovered by way of the zero restriction.

This methodology has been applied to compare racial and gender differentials in South African wages, though only with OHS data (Grün, 2004). In this paper, we attempt to move beyond the simple decomposition in order to identify the sources of unemployment. The longer time series we use also allows for a more accurate picture to enfold. It is important to note that the decomposition technique employed here ignores interactions between the seperate components, but already provides more information than cross section evidence is able to. Figure 1 compares two cross sections (October 1995 and September 2005). Unemployment probits are run on the set of age dummies, after which predicted probabilities of unemployment are obtained for each year. This picture suggests that the age-unemployment profile has remained largely unchanged, but that labour market circumstances have deteriorated uniformly for all age groups. The most credible way to separate age-specific and generational effects is by pseudo-panel decompositions. The decomposition results reported below suggest a flatter profile for younger individuals compared to the cross-section evidence, with more importance accorded to cohort changes in explaining the recent increase in unemployment.

Figure 1: Unemployment rate, by age, OHS 1995 and LFS 2005b



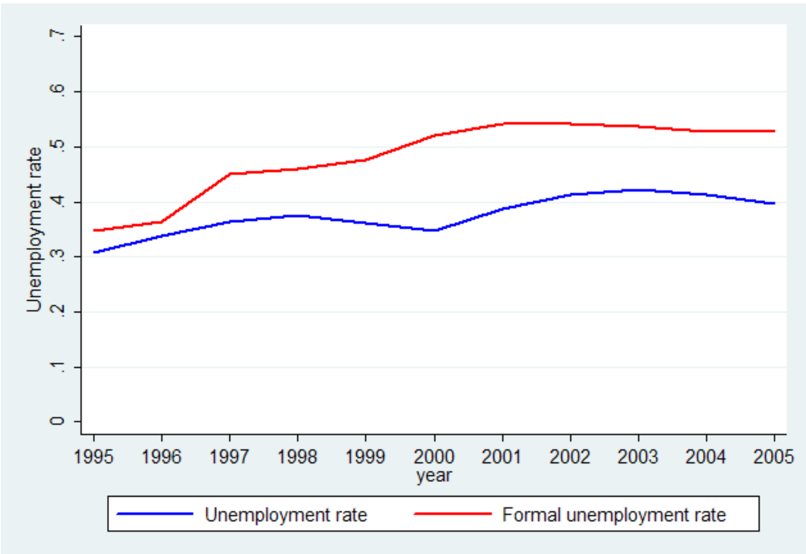
2.2 Data description

For this study we use all nationally representative South African household surveys that focused primarily on labour market issues and were conducted between 1995 and 2005: the 1995 to 1999 annual October Household Surveys, as well as the biannual Labour Force Surveys from 2000 to 2005.

The primary purpose of this study is to analyse changes in the unemployment rate. Kingdon and Knight (2006), by showing that the non-searching unemployed more closely represent discouraged work-seekers than the voluntarily unemployed, present convincing evidence that the broad definition of unemployment is a more accurate measure of the adequacy with which the economy is providing employment opportunities for the labour force. For the duration of this paper the unemployment and participation rates will therefore be calculated using the broad definition of the labour force.

A second issue to consider in the analysis of the unemployment rate at different time periods is the possibility that inconsistencies in sampling and questionnaire design will distort the true unemployment trend. This is primarily a concern for the earlier years in our sample, during which period Statistics South Africa (StatsSA) continually altered the questionnaires in order to improve the accuracy of the data. The effect of these changes is particularly evident in the improved capturing of informal economy workers and the large fluctuations in agricultural employment.

Figure 2: Unemployment rate and Formal unemployment rate, 1995-2005



Given the low wages and often unpleasant working conditions faced by informal sector workers (Casale and Posel 2002), it seems plausible to assume that most labour market participants would consider this as an employment option of last resort. Although the determinants of informal sector employment are of considerable importance in their own right, the interest of this study centers around the ability of the formal economy to provide employment for a rapidly expanding labour force. If poor formal sector job creation leads more people to resign themselves to a working life in the informal sector, we do not want our measure of unemployment to register this as a decrease in unemployment. For this reason we do not include informal sector workers amongst the employed in our calculation of the employment and unemployment rates. This also circumvents the problems associated with the inconsistent capturing of informal employment. It should be noted that since all informal sector workers are now counted as unemployed, our "formal sector unemployment rate" is higher than unemployment rates estimated in the conventional way. Our analysis will necessarily be silent on any issues that pertain to the informal economy.

Each of the five OHS's surveyed independently sampled households, but the LFS's were based on a rotating panel design, according to which only 20% of sampled households did not also appear in the previous survey. StatsSA has not released the LFS as a panel dataset and the coding of person identifiers do not facilitate the linkage of individuals across different waves of the LFS (Kingdon and Knight 2005: 18). There have been a few attempts at constructing panels from the different cross-sections, but it is too early to judge the reliability of the resulting analyses. Devey, Skinner and Valodia's (2006) work matches 5587 people across five of the LFS waves. However, in the absence of attrition and coding errors, the panel structure should have allowed the comparison of 20% of the 69150 working age individuals (or 13830 observations) that appeared in the February 2002 LFS, the first wave that they consider. This indicates that less than 40% of the original observations are apparently recoverable from the LFS's, which implies that studies using this methodology will suffer from a high degree of attrition bias unless this problem can be appropriately addressed.

Until September 2004, the LFS datasets were released with probability weights based on the 1996 Census, but the subsequent LFS probability weights were derived from the 2001 Census. In 2005 Stats SA re-weighted all the LFS's that originally used the 1996 Census according to the 2001 Census, in order to aid comparability across surveys. We make use of these re-weighted datasets in this paper. All the OHS's remain weighted according to the 1996 Census.

In creating the cohort panel, we average over individuals who share the same birth year. We construct a "birthyear" variable by subtracting *age* from the *year* in which the survey was conducted. This variable will suffer from some measurement error, since all individuals born between the day on which they were surveyed and the 31st of December will be assigned the birth year that actually follows their year of birth. This will be particularly severe in the March waves of the LFS's, where most individuals will actually be assigned the incorrect birth year. Given the fact that our empirical results show no discontinuities in the effect of birth year on unemployment, this is unlikely to pose a serious problem in our analysis.

The empirical analysis first proceeds without controls to trace pure cohort, age and year effects. Subsequently controls are introduced. Geographic heterogeneity is controlled for by including the averaged provincial dummy variables, which represent provincial shares for each cohort. We also wish to control for the variation in levels of education across cohorts. Since cohort employability can be affected by the distribution of education rather than just its mean, we opt for a flexible specification that allows for a differential effect of the different levels of education. The cohort sample means were constructed by averaging over four dummy variables that indicated whether each person had primary, incomplete secondary, complete secondary and any tertiary education. Individuals with NTC I, NTC II qualifications (or who held any certificate or diploma) but have not completed Grade 12 are considered to have incomplete secondary education. Individuals who held an NTC III qualification are counted as having complete secondary education.

An aggregated dummy variable which represents the proportion of over-age learners in a cohort is also included as a control. Shortly after the political transition, the Department of Education decided to normalise the age profile of learners in schools (Republic of South Africa, 1995, par 33). A part of this process entailed reducing the large numbers of over-age learners (defined in our study to be those older than 19) in school. Steps have been taken to find alternatives for this group in the form of adult education (Republic of South Africa, 1995, par 36). This stance could, however, partially explain surges in labour market participation, as many learners choose not to continue with further education in community learning centres. Furthermore, it is important to consider that the source of the over-age trend can be attributed to high repetition rates: should these individuals choose to exit the schooling system and enter the labour market, incomplete educational attainment reduces the employability of these candidates. In combination, these changes have potentially adverse impacts on unemployment levels in South Africa.

3 Decomposition of unemployment rate

3.1 Birth cohort panel decompositions

The empirical analysis commences by grouping all individuals born in the same year into cohorts, and applying the decomposition technique suggested by Deaton (1997) to the cohort averages of the unemployment rate. This choice of aggregation delivers 561 cohorts, with an average of 1917 sampled working-age individuals per cohort. The largest of these consists of 5277 and the smallest of 289 observations. Unfortunately there are 22 cohorts that contain less than 100 sampled labour force *participants* – the acceptable threshold required to ignore sampling errors (as in section 2.1) – from which we can calculate our unemployment rates. Since the average number of labour force participants per cohort is 1205, only a small number of cohorts will suffer from non-negligible inaccuracies attributable to sampling errors. Therefore the inconsistency that arises from using conventional fixed effects estimators is unlikely to play an important role in our results.

In section 3.1.1 we calculate the average unemployment rate over all individuals in a birth cohort, and regress this on the fifty age dummies representing the ages of 16 to 65 (age 15 is chosen as the reference age group), on the sixty birth cohort dummies, which represent being born in the years 1931 to 1990 (1930 is the reference birth year) and on the nine transformed year dummies. In section 3.1.2 the same method is applied to the participation and employment rates. Since both of these averages are calculated using all individuals of working age (as opposed to only labour market participants), all cohorts consist of more than 100 observations.

3.1.1 Decomposition of unemployment rate by age, cohort and year

Figure 3 shows the decomposition of the unemployment rate into age, cohort and year effects. The lines in the graph of figure 3.1 depict the unemployment rates experienced by every third birth cohort at different ages – showing each cohort produces a cluttered graph. For example, the leftmost line represents the unemployment rates for the birth cohort aged 15 in 1995 (and hence born in 1980) for all the years from 1995 and 2005. This curve is positioned above that of the next youngest cohort until the age of 23, after which the two curves intersect: this implies that the younger cohort faced higher unemployment than individuals born three years earlier at the same ages. Since most of the lines are above those directly to their right, this reveals that individuals born more recently generally experience higher unemployment rates than older birth cohorts. The only reason for an overlap in the curves is that unemployment rates for most cohorts show a decline in the last few years of our sample. It is possible to express the same information in three dimensions, as in figure A1 in the appendix.

The year effects (in figure 3.4) represent the impact of macroeconomic events on the unemployment rate. Since the separate year effects add up to zero by design, it captures the business cycle variation in the unemployment rate. It can be observed that between 1995 and 1997 there was a steep increase in cyclical unemployment, followed by a brief decline until 1999, before reaching its highest year effect in 2000. After this, the unemployment rate showed a steady decline. The magnitude of the increase in the cyclical unemployment effects between 1995 and 1997 is primarily driven by an implausibly large decrease in agricultural employment reported in the OHS's over this period. The South African Reserve Bank identified the third quarter of 1999 as the start of an upswing phase in the South African economy (SARB, 2006: S159). It is interesting to note that the cyclical variation in unemployment shows some correspondence to the business cycle, but appears to lag the cycle by about a year.

The age profile of unemployment (figure 3.2) differs markedly from that presented in Figure 1, which demonstrates the value of using the cohort panels rather than cross-sectional analyses. Age appears to be an unimportant determinant of unemployment between the ages of 15 and 40, but unemployment increases rapidly amongst labour force participants of older ages.

The birth cohort graph (figure 3.3) reveals that generational effects played the most important role in the increase in unemployment (as judged by the size of the cohort coefficients, labeled on the y-axis). In many respects, the cohort effects represent the most important of the three composite parts, as it is a reflection the longer-term trend in the economy. Changes in the birth cohort effect can result from structural changes at the macro-economic level, shifts in the preferences of individuals or differences in the productive characteristics (observed or unobserved) across generations. In this case, the cohort effects show a deterioration of the labour market prospects of younger generations, something which is of obvious concern for the growth outlook of the economy. The cohort effects combined with the age profile explains the U-shaped age-unemployment profile observed in cross-sections: the high unemployment experienced by the young is due to the disadvantage they face as a result of entering the labour market in a period of higher unemployment, rather than the inherent fact that they are young. In contrast, the higher unemployment rates amongst older individuals, is explained by their age and occurs despite profiting from the lower unemployment-cohort effect. This separation of cohort and age effects is not possible without the additional information provided by the pseudo-panel.

It is important to emphasise that the decomposition technique used here does not allow us to make any causal inferences on the determinants of unemployment. It does not, for instance, tell us why age is positively correlated with unemployment at ages older than 40. It merely serves to separate out the covariation between unemployment and age, birth cohorts and years. In section 3.3, our empirical investigation will take us beyond the simple decomposition employed here in controlling for a set of productive and demographic characteristics.

Figure 3: Unemployment rate by cohort and their decompositions, 1995-2005

Figure 3.1 Unemployment rate by birth cohort and age

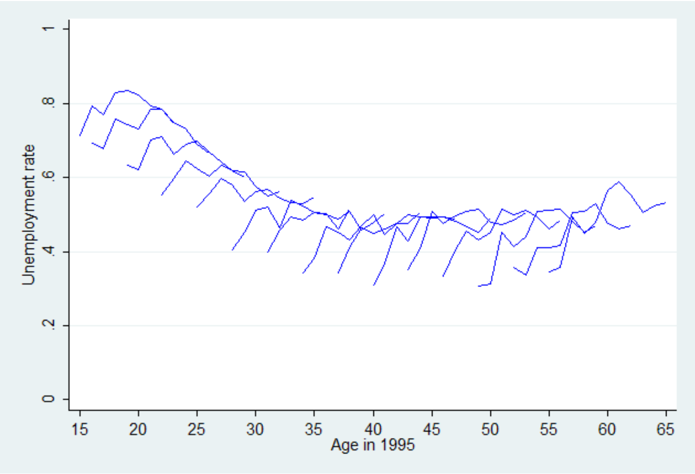


Figure 3.3 Unemployment rate birth cohort effects

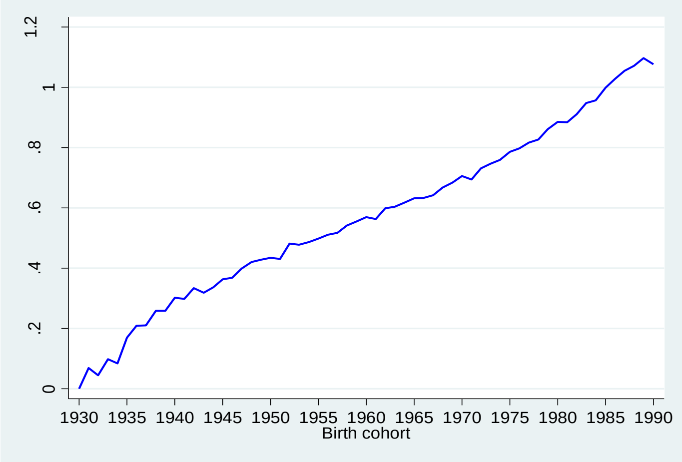


Figure 3.2 Unemployment rate age effects

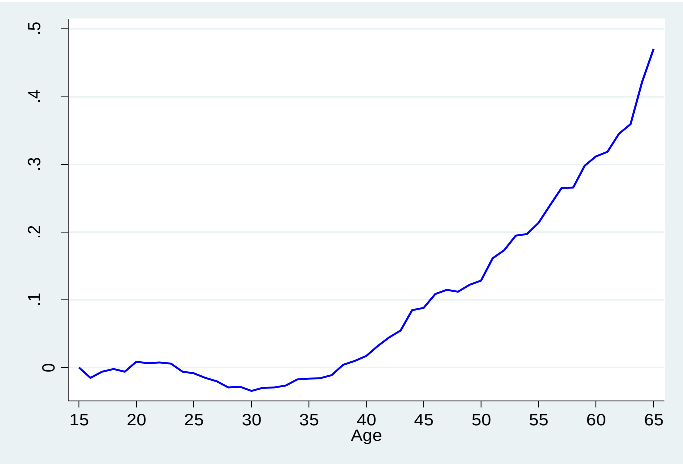
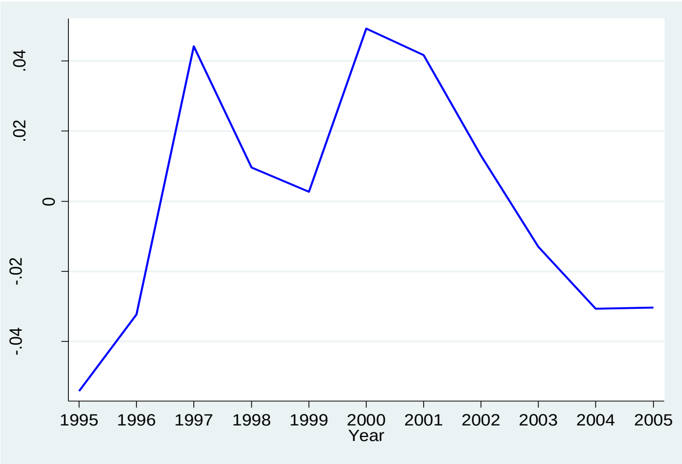


Figure 3.4 Unemployment rate year effects



3.1.2 Decomposition of employment and participation rates by age, cohort and year

It was stated in section 2.1 that the unemployment rate is equal to one minus the employment rate divided by the participation rate. In order to distinguish between the effects of these composite factors of unemployment, the same decomposition technique will now also be applied to the employment and participation rates separately. The raw employment and participation rates are presented by age and birth year in figure 4.1 and 5.1 (and by birth year and year in figures A2 and A3).

Figure 4.4 indicates that the large increase in the unemployment year effects between 1995 and 1997 was driven by the rapid decrease in the cyclical component of the employment rate. Again this is mainly attributable to the inconsistent sampling methodology referred to in section 3.2: over this period the OHS's indicate an improbably large decrease in agricultural employment. The increase in unemployment occurred despite the fact that the cyclical component of labour force participation showed a modest decline (figure 5.4). Between 1997 and 2005, the employment rate year effects showed a steady resurgence, which dominated the increase in participation rates between 1997 and 1999, but not in 2000. After 2000, a stable employment year effect and decreasing participation combined for a decrease in the cyclical component of the unemployment rate.

Looking at the age profiles (figures 4.2 and 5.2), we observe that both employment and participation are characterised by an inverted U-curve. The similarity in the shapes of the age effects of participation and employment for people younger than 40 explains why age does not appear to play an important role in determining unemployment: increases in participation are matched by increases in employment, so as to cancel each other's effect on the unemployment rate. After the age of 40, the employment rate starts to drop rapidly, whereas the participation rate is marked by a more steadily decline. This means that older labour force participants are "losing jobs" faster than they are leaving the work force, which exerts a positive net effect on the unemployment rate.

The birth cohort effects for employment and participation reveal that individuals from younger generations have higher participation rates and lower employment rates than those born earlier, both of which imply that unemployment will be higher for this group. From Figures 4.3 and 5.3 it is clear that the participation cohort effects show larger changes in coefficient magnitudes than the decrease in the employment cohort effects. The non-linear manner in which these two factors combine to determine unemployment makes it difficult to gauge the relative contribution of each of these effects by comparing their coefficients. An increase in the participation rate will increase the unemployment rate by approximately $\frac{e}{p^2}\ p$ (where e and p denote the employment and participation rates), whereas a decrease in the employment rate will increase the unemployment rate by $\frac{1}{p}\ e$. At full employment, the effect of these changes will be the same, but for any positive level of unemployment the unemployment rate will respond more to a decrease in the employment rate than to an increase in the participation rate. Evaluating these equations at the sample means of e and p (39% and 68%) shows that a change in the employment rate will have an effect on the unemployment rate that is about 60% larger than a change in the participation rate of the same size (in percentage points). At the average participation rate, the observed decrease in employment faced by the different birth cohorts would have led to a 19 percentage point differential between the unemployment rates for the youngest and oldest birth cohorts (abstracting from age and year effects), whereas the observed increase in participation rates would have (at the mean employment rate) increased the unemployment rate by 80 percentage points. It is therefore the case that approximately 81% of the increase in unemployment rates faced by the youngest generation was caused by their increased labour force participation rates, whereas the lower employment rate contributed the remaining 19%.

3.2 Birth cohort and population group panel decompositions

In this section we further disaggregate our cohorts by population group. To simplify our analysis, and to avoid problems associated with small cohort sizes, the focus will fall on comparing the black and white population groups only. This gives us 1122 cohorts, with an average of 1462 observations for black and 167 observations for white cohorts. The largest cohort consists of 4336 and the smallest cohort of 38 observations. As above, for the unemployment decomposition, cohort sizes are only determined by active labour market participants. 300 cohorts in the dataset are constructed from less than 100 labour market participants, 264 of which are from the white population group. The potential for inconsistency in our estimators arising from measurement error therefore increases substantially when moving from a birth cohort panel to a birth year and population group cohort panel.

In section 3.1 our interest lay only with the shapes and relative contributions of the different components of unemployment (as well as employment and participation). In comparing the age and birth cohort effects of blacks and whites, we are now also interested in the level of these curves and therefore the graphs in Figures 6 to 8 were drawn to include the effect of the constants from the decomposition regressions. The fact that the year effects are restricted to add up to zero means that the absolute levels of these curves carry no meaning: hence, the year effect graphs omitted the effect of the constants.

Figure 4: Employment rate by cohort and their decompositions, 1995-2005

Figure 4.1 Employment rate by birth cohort and age

Figure 4.3 Employment rate birth cohort effects

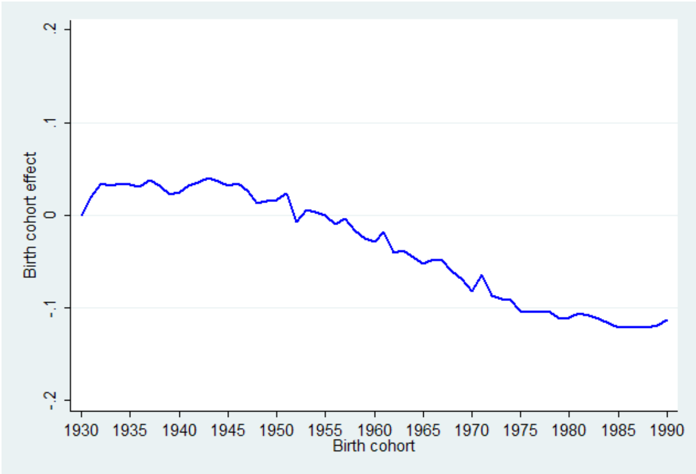


Figure 4.2 Employment rate age effects

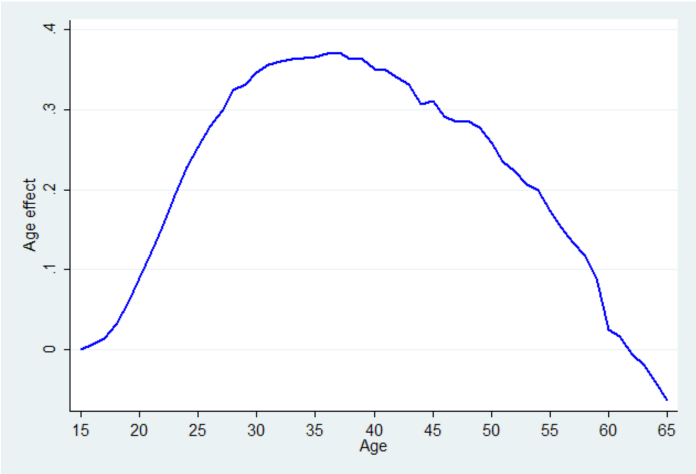


Figure 4.4 Employment rate year effects

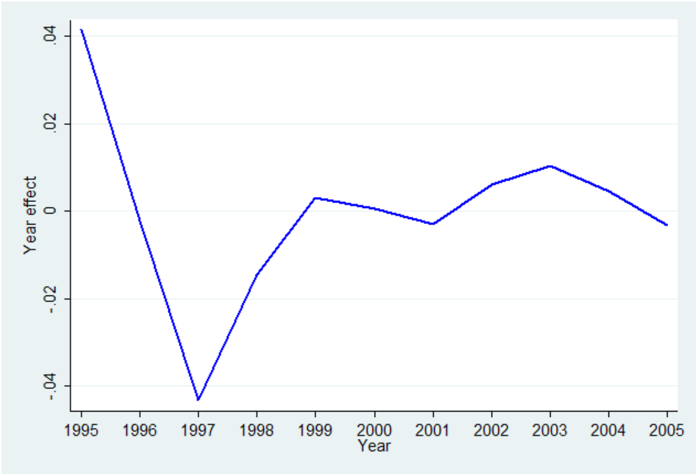


Figure 5: Participation rate by cohort and their decompositions, 1995-2005

Figure 5.1 Participation rate by birth cohort and age

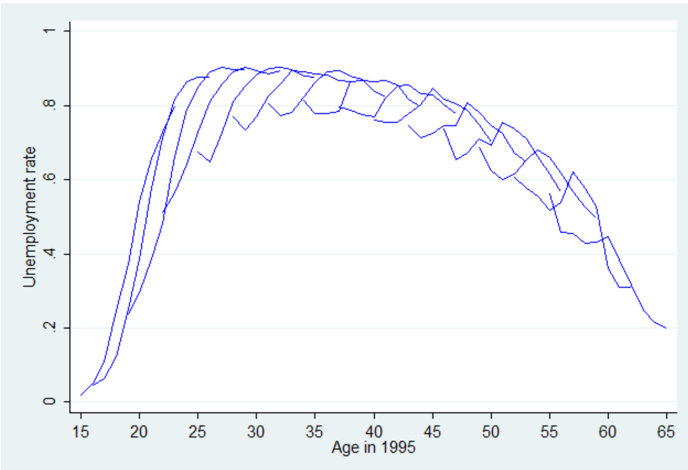


Figure 5.3 Participation rate birth cohort effects

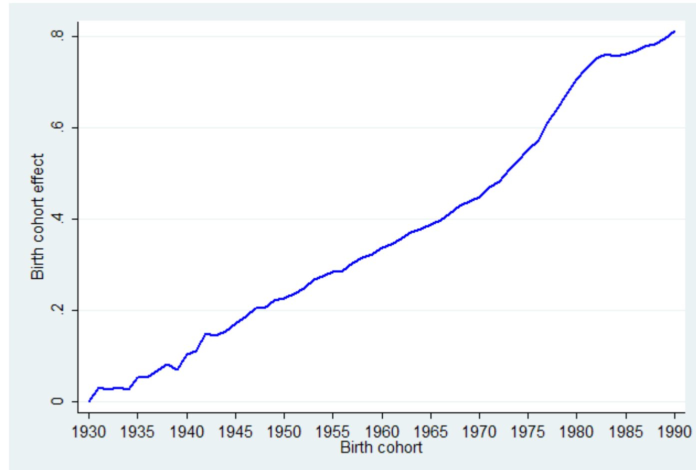


Figure 5.2 Participation rate age effects

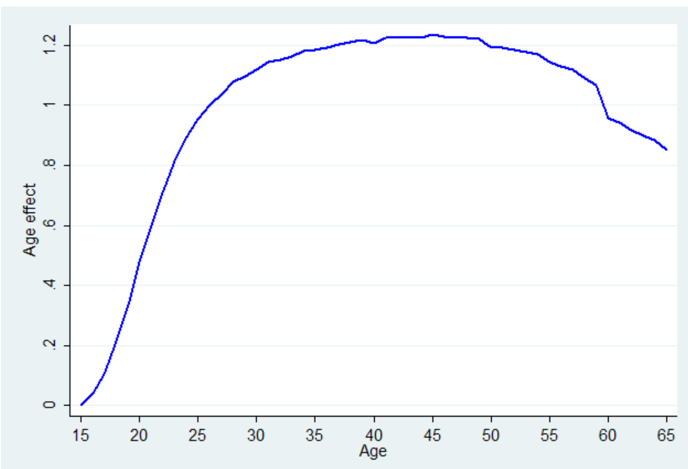
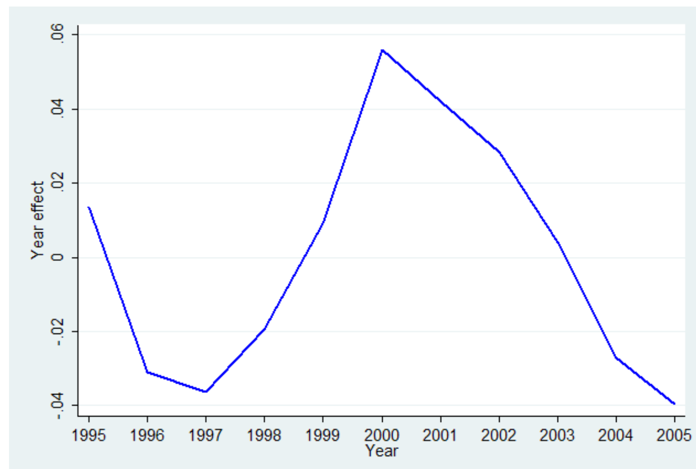


Figure 5.4 Participation rate year effects



3.2.1 Decomposition of unemployment rate by age, cohort and year

Figure 6 shows the decomposition for the unemployment rates by birth cohort and population group. From the raw unemployment rates, it can be observed that unemployment is higher for blacks than for whites of all ages and birth cohorts. The figure suggests that youth unemployment is more of a problem, and persists until older ages for black workers relative to their white counterparts, although white unemployment shows a sudden spike (and a high degree of volatility) at ages younger than 25. The higher unemployment experienced by older labour force participants is a phenomenon that appears to be restricted to the black population.

The year effects for the black and white unemployment decompositions reveal a large degree of correspondence, except that white unemployment experienced a decrease between 1995 and 1996, compared to an increase for the black population. The reported decrease in agricultural employment over this period fell disproportionately on blacks, due to their greater (relative) representation in this sector.

The unemployment age effects displayed in figure 6.2 differ markedly between the two population groups. The black age profile is similar to that of the total population (figure 3.2), except that unemployment shows a small decrease between the ages of 20 and 30 before increasing rapidly after the age of 40. The white age-unemployment curve also show a decrease in unemployment between the ages of 20 and 30, but the subsequent increase is much flatter than what we observe for the black population.

In figure 6.3 it can be observed that black labour market participants from younger generations face higher unemployment rates, which is a trend that is consistent with the cohort effects observed for the labour force as a whole. In comparing the black and white unemployment birth cohort effects, we see that blacks and whites born before 1940 face similar unemployment cohort effects. For slightly younger birth cohorts the black workers start to experience higher unemployment cohort effects than that of whites, and this disadvantage grows as we move to more recent birth years. Amongst the very youngest birth cohorts, a sharp increase in white unemployment cohort effects can be observed, although the sudden increase in between-birth year volatility (in contrast to the relatively smooth changes observed elsewhere) is indicative of estimator inaccuracies, driven by the small number of white labour market participants sampled from these birth cohorts.

3.2.2 Decomposition of employment and participation rates by age, cohort and year

A comparison of the black and white participation rates by cohort (figure 8.1) shows that black participation rates were initially much lower than that of whites (especially amongst the young), but that a convergence took place between 1995 and 2005. This sharply contrasts with the large and persistent differences in employment rates shown in figure 7.1. The absence of a large increase in black employment over the period implies that the increase in participation rates must have resulted in rising unemployment rates, and that this trend was particularly strong amongst young blacks. The previous section showed that this was indeed what we observed.

The decomposition shows that the year effects' impact on employment were similar for the black and white population groups, except for the much larger decrease in employment experienced by blacks between 1995 and 1996. The cyclical component of the black and white participation rates also appear to move in unison, although the black year effects show larger fluctuations than that of whites.

The age profiles of participation follow an inverted U-shape, and are very similar for the two races. Participation rates are higher amongst young whites than blacks, but the more rapid increase for blacks at young ages means that from the age of 30 onwards the age profiles are nearly identical. The employment age profiles show that white youths experience a rapid increase in their probability of being employed as they move from the age of 18 to 25, after which the employment rate remains high until steadily decreasing after the age of 50. Black youths experience a much slower increase in their employment likelihood and the age-employment profile starts to decrease around an age of 40. This explains the U-shaped form of the black age-unemployment curve, as opposed to the relatively flat profile of whites.

The birth cohort effects of the participation rate indicate that younger generations tend to have a higher proclivity towards labour force participation. This increase is sharper for blacks, who experience lower participation effects for older cohorts, but similar effects for the younger cohorts. The birth cohort profiles of the employment rate for the two population groups, on the other hand, show very different trends, with the white employment rate increasing for more recent birth years and black employment decreasing for younger cohorts. The increases in white employment therefore moved in the same direction as labour force participation, whereas the stronger increase in black participation rates was met with a *decrease* in employment. This explains why the white unemployment cohort effects remained more or less constant, as opposed to the large increase in the unemployment rate for younger black birth cohorts. The formal economy has clearly not been able to provide jobs for the rapidly expanding labour force, and the burden of this failure has fallen disproportionately on black youths.

The decrease in the employment birth cohort effects suggests that black unemployment would have increased even in the absence of increasing participation rates. Using the same linearisation as in section 3.1.2, we find that 71% of the generational increase in black unemployment, and 56% of the white increase, was driven by an increase in labour force participation, as opposed to a decrease in employment.

Figure 6: Unemployment rate by cohort and their decompositions, 1995-2005

Figure 6.1 Unemployment rate by birth cohort and age

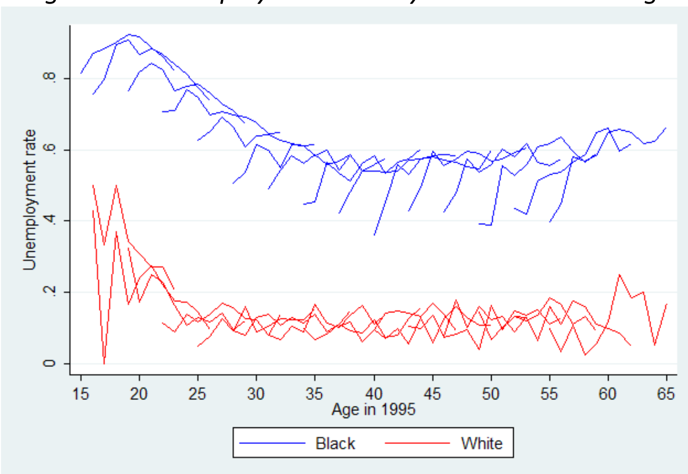


Figure 6.3 Unemployment rate birth cohort effects

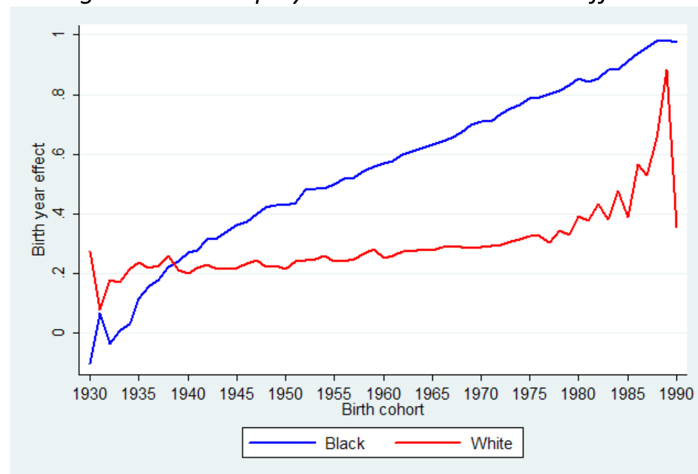


Figure 6.2 Unemployment rate age effects

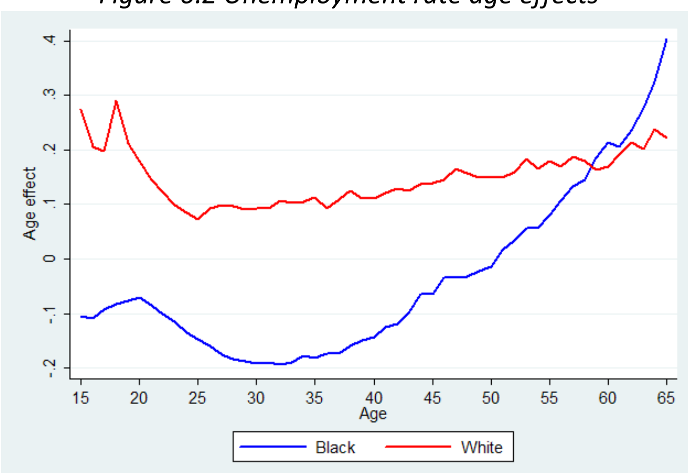


Figure 6.4 Unemployment rate year effects

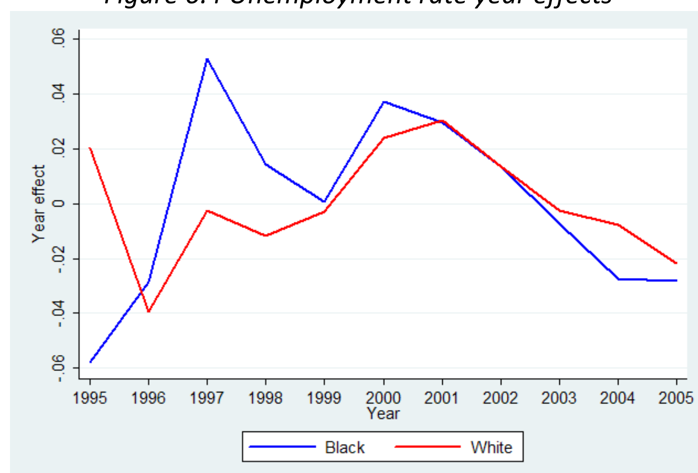


Figure 7: Employment rate by cohort and their decompositions, 1995-2005

Figure 7.1 Employment rate by birth cohort and age

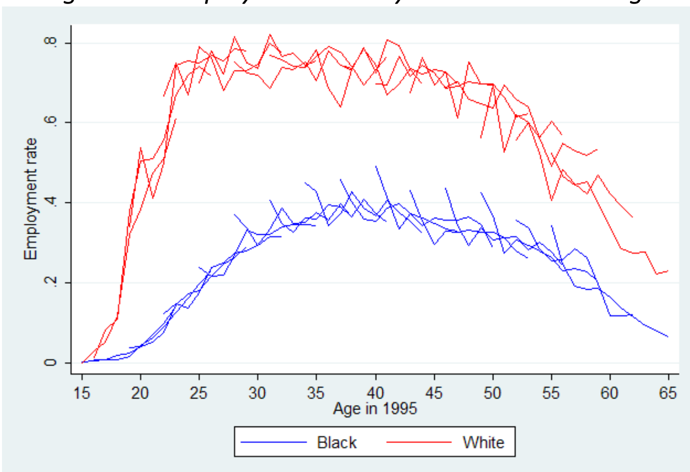


Figure 7.3 Employment rate birth cohort effects

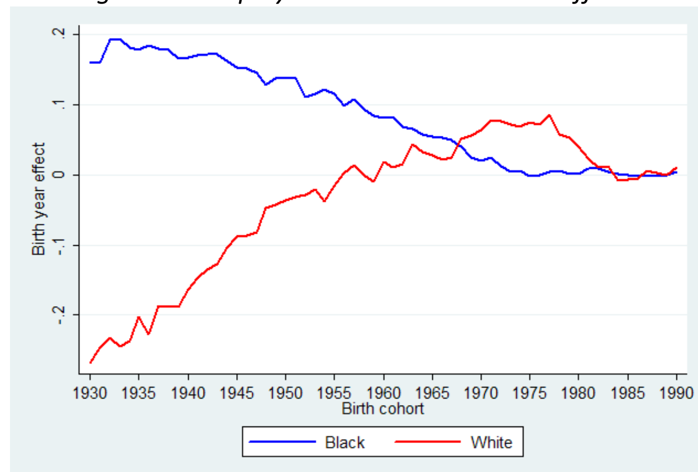


Figure 7.2 Employment rate age effects

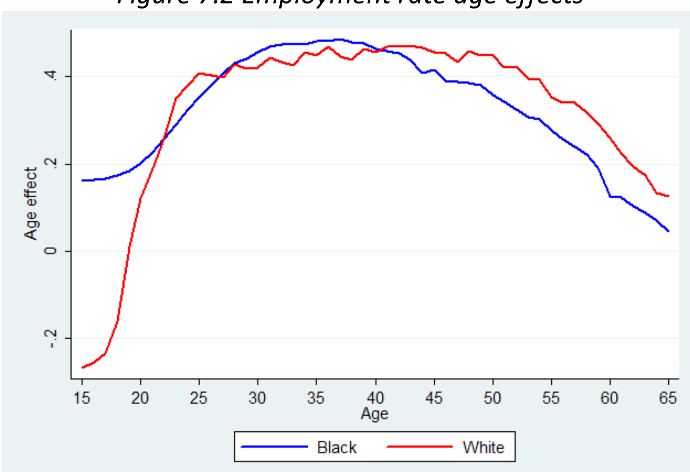


Figure 7.4 Employment rate year effects

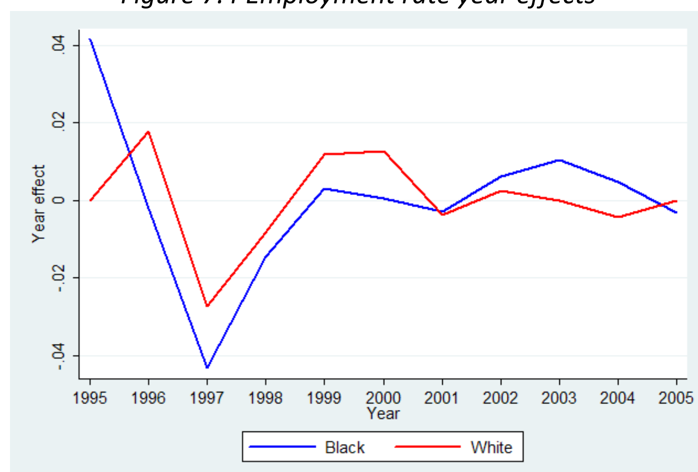


Figure 8: Participation rate by cohort and their decompositions, 1995-2005

Figure 8.1 Participation rate by birth cohort and age

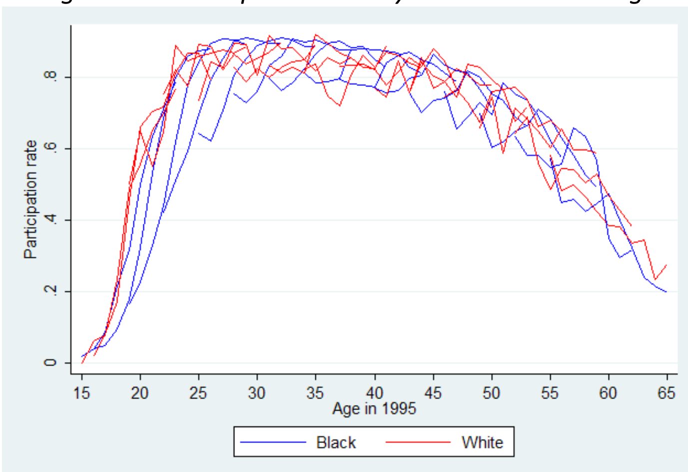


Figure 8.3 Participation rate birth cohort effects

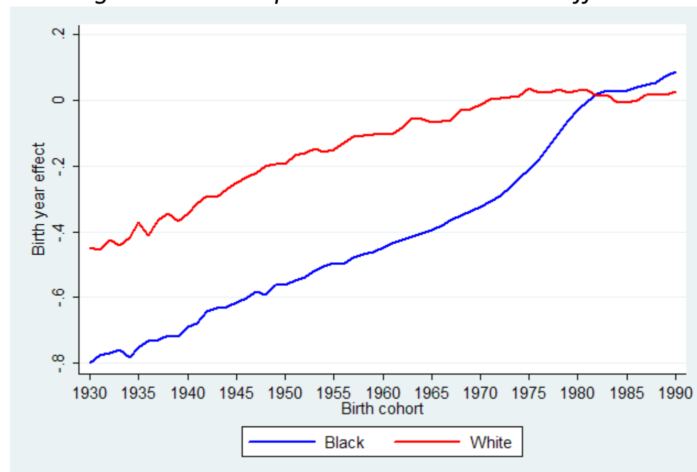


Figure 8.2 Participation rate age effects

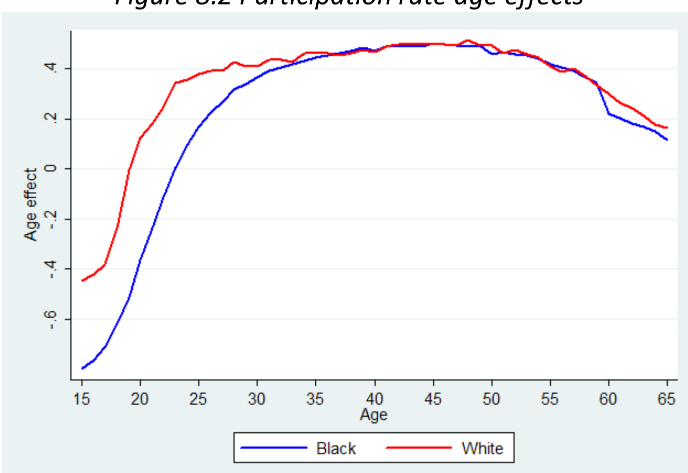
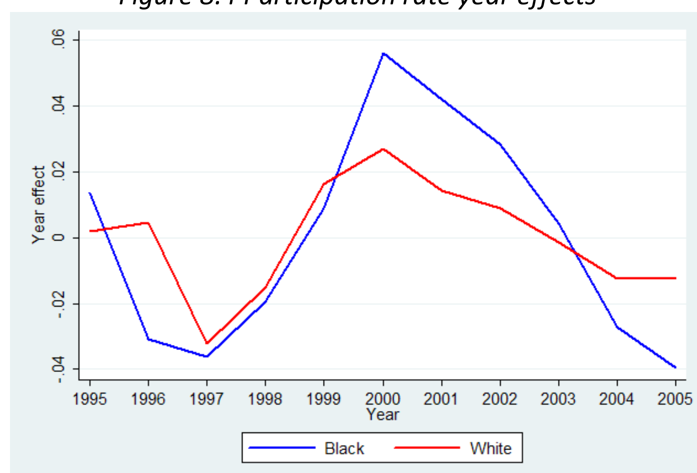


Figure 8.4 Participation rate year effects



3.3 Decompositions with controls

The above analysis reveals that the labour market transitions of the black groups most closely represent the movements of the overall working age population. This is expected, given that this is the largest population group within South Africa. It can be observed, however, that the white group has not followed the same trends. Can these differences be explained by changes in the average level of observable characteristics? The cohort effects of the different sub-populations do not only capture the differential impact of long-term macroeconomic and productivity movements, but also the discriminatory features of the labour market. The latter appear in different forms: differences in educational quality (both between population groups, and over time) has long been entrenched by separately operated education departments; some cohorts may also have sub-standard educational *attainment* which can be attributed to various circumstances (such as political unrest, poverty and other demographic features). A further source is taste discrimination the extent of which can be ascertained most concretely by controlling as exhaustively as possible for observable characteristics. The section which follows undertakes a descriptive analysis to identify potential sources of differential unemployment realisations. Section 3.3.2 augments the preceding empirical arguments, by introducing additional controls to the decompositions. It is assumed that these productive and demographic characteristics are correlated with the cohort fixed effects, and hence the least squares dummy variable model (or fixed effects estimation) remains the most suitable channel to continue with the analysis.

3.3.1 Descriptive analysis

Figure 9 highlights differential educational attainment, both across generations and population groups. The sharp changes observed for the most recent birth cohorts should be ignored, as many of these group members are still in transition from one education level to another. What is revealing, however, is the changing face of black education. The oldest generations were very unlikely to move beyond primary education, while fewer still progressed to matriculation or tertiary qualifications. For younger cohorts, a sharp decline in primary attainment is accompanied by modest transitions to some high school education, with somewhat larger probabilities of completing secondary schooling and obtaining tertiary education. For all cohorts, most whites have moved beyond primary education. While many members of older cohorts did not complete secondary education, this is rarely the case for their younger counterparts, many of whom also move on to the tertiary level. Human capital theory therefore suggests that the younger generations of both racial groups should be more productive, and hence face *less* difficulty in becoming employed. This is not what is observed in the cohort profiles of the preceding analysis. What can possibly be accounted for, is the racial differential in unemployment.

Figure 9: Educational attainment, by birth year and population group

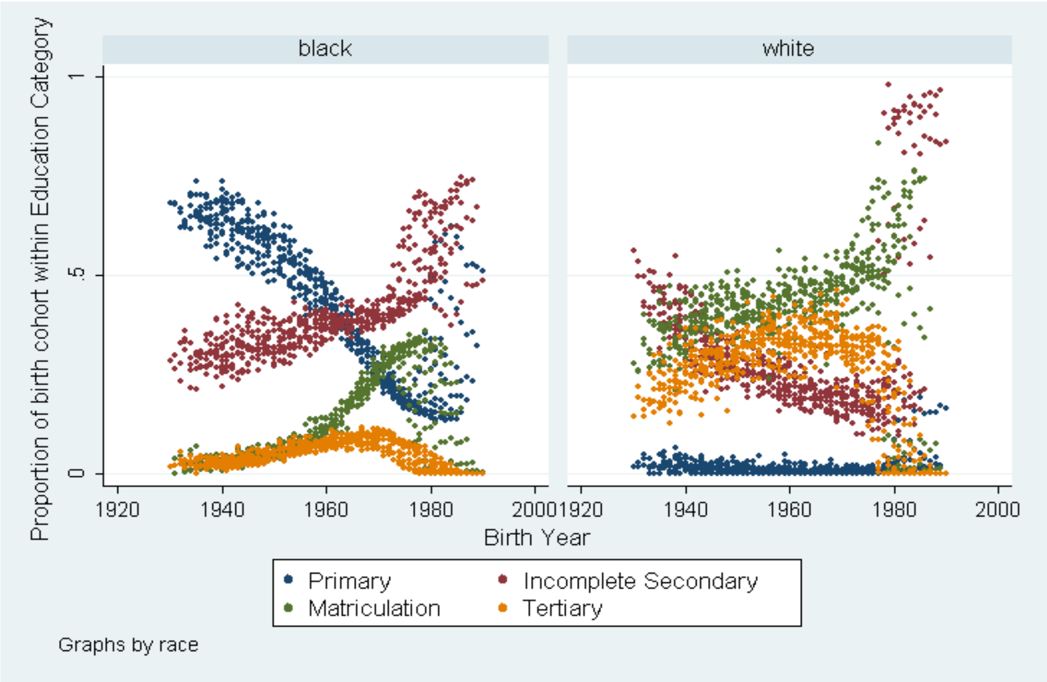


Figure 10: Proportion of over-aged learners, by year and population group

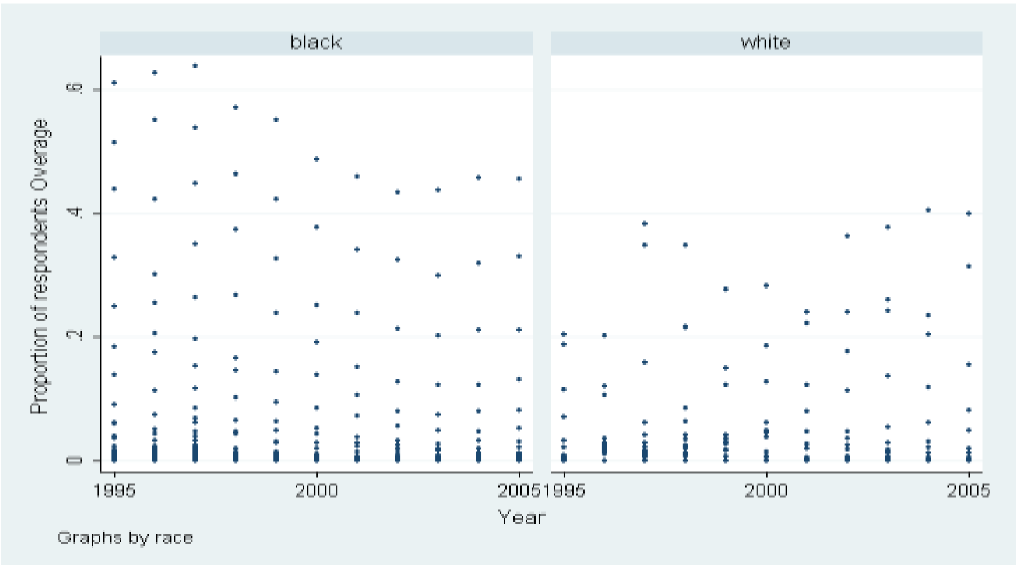


Figure 10 investigates the effects of over-aged learners. It is clear that black learners are more likely to still be in the schooling system after the age of 19 than white learners, but that the prevalence thereof has been declining. A change in the Department of Education's policies could therefore have had a substantial effect on the labour market: by not allowing individuals to attend school, they could have been "forced" into the labour market earlier than would have otherwise been the case, hence increasing the labour force participation rate and possibly also the unemployment rate. From figure 10 we would expect this to have had a more pronounced effect on the black cohorts.

Figure 11: Proportion of individuals living in rural areas, by age and population group

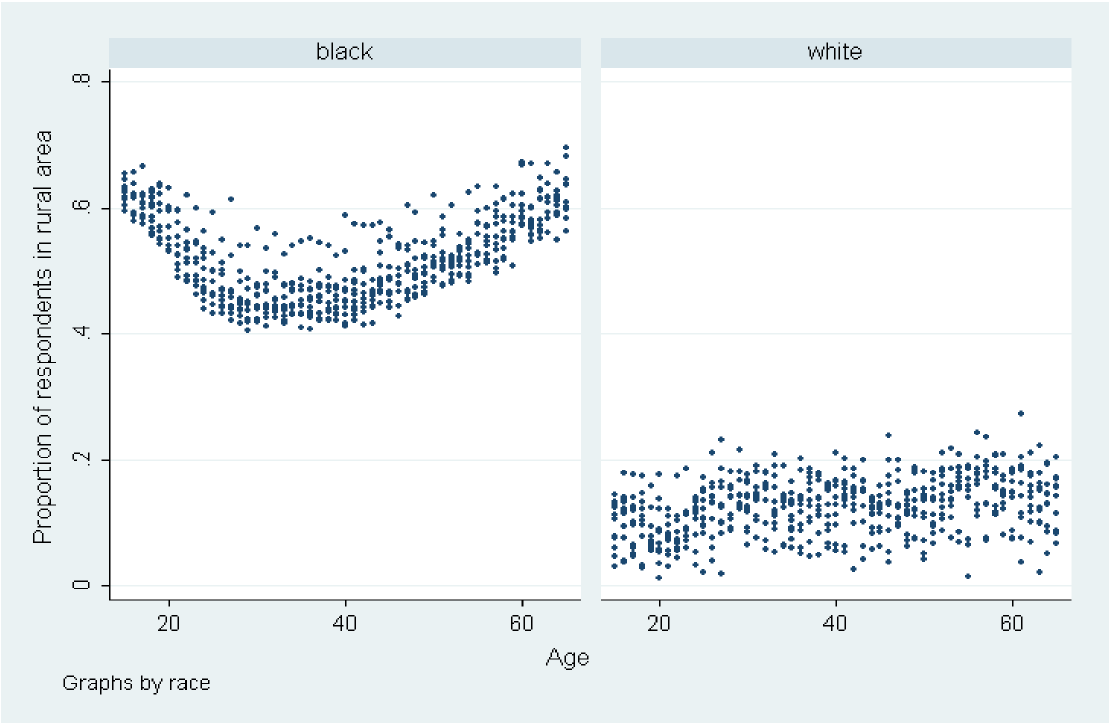


Figure 11 shows the tendency of different age groups to reside (and consequently seek work) in rural areas. Many studies have found an effect of area of residence (rural or urban) on the probability of being unemployed, which might be explained by the disproportionate number of younger blacks choosing to stay and supply labour in these regions. A definite reduction in this rate for middle-aged individuals highlights the tendency to migrate to cities to find employment. It seems as if many older individuals subsequently return to their rural homes, while the youngest individuals have not yet embarked on the urban route. This could induce a glut in the youth labour market, with participants only seeking urban employment opportunities at a slightly later age. Whites are substantially more urbanised, and uniformly so at all ages. The latter observation explains possible interracial differences, but within racial variation (by age) is unlikely to account for much of the increase in white unemployment.

3.3.2 Control Variables

Table 1 presents the magnitudes and significance of only the variables used to control for racial and generational differences. A brief look at figures 12 to 14 reveals that the inclusion of explanatory variables does not alter the shapes of either the cohort, age or cyclical profiles dramatically. Modest gradient adjustments occur, nevertheless. This indicates that differences in geographical location, household composition and even educational attainment does not fully account for racial or generational differences.

The presence of non-linear effects of education on labour market outcomes is clearly visible. High shares of tertiary education reduce unemployment rates within black cohorts, as does complete secondary education albeit less successfully. Incomplete secondary education is only slightly more favourable than primary education in this regard, but only the effect of tertiary education differs significantly from primary education for black cohorts. In the case of the employment rate, a progression from low employability to high employability also results from higher attainment. This carries over to participants, who presumably anticipate that their higher levels of human capital will be rewarded in the labour market. One notable exception in this picture is the high, positively significant coefficient on complete secondary for black participation, which suggests that matriculation is associated with too high a degree of optimism. For whites, the coefficients on educational attainment are largely insignificant, bar for the important effects of tertiary education on participation and employment.

Provincial effects are only significant in determining black employment and participation, with high penalties for living in the Eastern Cape, Mpumalanga and KwaZulu-Natal. These provinces have a large poor rural contingent, and may not have the capacity to generate large numbers of formal sector jobs. It is interesting to note that participation tracks employment in magnitude: this shows that potential participants rationally gauge their prospects in the labour market by current employment levels, and choose to participate accordingly. For this reason no strong (statistically significant) provincial differences appear in unemployment rates.

Variables that explain household composition are significant in almost all cases. Cohorts with high concentrations of household heads experience less unemployment: high participation rates and even higher positive employment prospects are associated with these groups. Marriage reduces unemployment for black cohorts, which is constituted by higher employment rates. Participation rates are lower in cohorts with high marriage rates, although it may be more instructive to distinguish this effect by gender. The number of unemployed household members is particularly significant for the white cohorts: this scenario exacerbates unemployment vulnerability by way of lower employment probabilities. At the same time, other household members enter the labour market to strengthen the household safety net.

The prevalence of over-aged students has ambiguous effects on unemployment. It however strongly reduces participation, as these individuals remain out of the labour market for longer periods due to grade repetition. It reduces white employment and increases black employment prospects.

Introducing the lag of the log of wage earnings has the effect of adjusting all the profiles substantially, with some equalisation occurring both across generations and racial groups. It is interesting to note that this variable was insignificant in all regressions, but had the largest ability to absorb cohort effects. Unfortunately, the evidence suggested that lagging the wage did not solve the potential simultaneity between unemployment and earnings, so that this variable was omitted from our final specification.

Table 1: Control variable coefficients

	Unemployment		Employment		Participation	
	Black	White	Black	White	Black	White
Incomplete Secondary	-0.014 (0.906)	0.348 (0.697)	-0.026 (0.622)	0.13 (0.418)	-0.064 (0.342)	0.224 (0.181)
Complete Secondary	-0.13 (0.389)	0.299 (0.734)	0.125 (0.108)	0.199 (0.221)	0.37 (0.000)***	0.275 (0.121)
Tertiary	-0.349 (0.024)**	0.223 (0.803)	0.414 (0.000)***	0.276 (0.100)*	0.128 (0.257)	0.32 (0.075)*
EC	0.113 (0.650)	0.037 (0.749)	-0.517 (0.001)***	0.095 (0.227)	-0.33 (0.078)*	0.074 (0.365)
NC	-0.175 (0.642)	-0.116 (0.388)	-0.152 (0.440)	0.067 (0.453)	-0.143 (0.564)	0.012 (0.903)
FS	0.146 (0.556)	-0.099 (0.425)	-0.431 (0.006)***	0.036 (0.643)	-0.255 (0.177)	0.1 (0.231)
KZN	-0.03 (0.912)	-0.077 (0.591)	-0.476 (0.001)***	-0.055 (0.498)	-0.339 (0.054)*	-0.041 (0.622)
NW	-0.015 (0.957)	-0.161 (0.302)	-0.499 (0.001)***	0.093 (0.302)	-0.286 (0.141)	0.124 (0.212)
GAU	0.17 (0.480)	-0.18 (0.127)	-0.471 (0.001)***	-0.002 (0.966)	-0.277 (0.145)	0 (0.997)
MPU	0.321 (0.248)	-0.062 (0.692)	-0.454 (0.004)***	-0.144 (0.125)	-0.409 (0.025)**	-0.111 (0.276)
LIM	0.242 (0.351)	-0.035 (0.854)	-0.384 (0.010)***	0.029 (0.800)	-0.262 (0.158)	0.067 (0.597)
Married	-0.452 (0.000)***	-0.126 (0.104)	0.138 (0.003)***	0.053 (0.348)	-0.113 (0.052)*	0.029 (0.632)
Household Head	-0.197 (0.038)**	-0.183 (0.003)***	0.297 (0.000)***	0.265 (0.000)***	0.401 (0.000)***	0.211 (0.000)***
Number of Unemployed in Household	0.055 (0.300)	0.376 (0.000)***	-0.068 (0.006)***	-0.181 (0.001)***	0.047 (0.187)	0.113 (0.035)**
Overage	-0.111 (0.159)	0.329 (0.054)*	0.112 (0.051)*	-0.776 (0.000)***	-0.482 (0.000)***	-0.666 (0.000)***
Constant	-0.129 (0.612)	0.195 (0.786)	0.669 (0.000)***	-0.366 (0.022)**	0.022 (0.907)	-0.584 (0.001)***
Observations	561	560	561	561	561	561
R-squared	0.97	0.768	0.989	0.982	0.995	0.984

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

3.3.3 Effect of controls on differentials

Figures 12 to 14 form the basis of the following discussion. Each panel shows the original decompositions along with the profiles once the controls (as above) are introduced. Cohort and age profiles are again adjusted for their respective constants to facilitate interpretation. However, the approach is different for controlled and uncontrolled graphs. Compare the two regressions, the first of which represents the simple decomposition, and the second which introduces the controls:

$$y = c_1 + x_1'\beta_{age\,1} + x_2'\beta_{cohort\,1} + x_3'\beta_{year\,1} + \varepsilon_1$$

$$y = c_2 + z'\beta_{controls} + x_1'\beta_{age\,2} + x_2'\beta_{cohort\,2} + x_3'\beta_{year\,2} + \varepsilon_2$$

The x_i $(i = 1 \dots 3)$ are the dummy variable vectors, while z is the vector of controls introduced subsequently. It is clear that the constants, c_1 and c_2 might differ substantially, and suggest a level shift in the profiles. Since we are concerned with the changes of β_i*'s* effects, it is evident that the controlled profiles should not just be scaled by c_2 but by $c_2 + z^{'}\beta_{controls}$. To this end, it would be required that a mean value of z be inserted to determine the shift paramenter. We therefore choose to use a single cohort's mean characteristics (those of birthyear 1990 for each respective population group) to execute the adjustment.

It is evident (in Figures 12.4, 13.4 and 14.4) that in each case year effects are not dramatically altered by controls. This shows that the cyclical variation in labour market outcomes (as well as survey-specific sampling error) is not correlated with educational attainment and demographic factors. This is to be expected, since individual characteristics should not vary over the business cycle.

For unemployment, the cohort profiles do not undergo any notable changes in either shape or gradient (figure 12.3). Therefore educational attainment does not fully control for generational differences in unemployment (as expected in the descriptive analysis). Age effects are largely unchanged for whites (figure 12.2), except for older individuals, who experience an upward shift. For blacks, the U-shaped profile disappears, with a strong upward age trend in unemployment. This is largely due to implicitly controlling for the impact of rural residence by the provincial variables.

Figure 13.3 reveals that controls again do not account for any substantial generational or racial differences in employment. Only mild shifts occur, though the endpoints of the profiles correspond strongly to their uncontrolled counterparts. The age profiles are of greater interest. Youth employment rates remain unaffected by controls, though employment levels of older age groups are somewhat subdued by the conditional variables. This explains the greater unemployment effects witnessed at older ages.

Figure 14.3 reveals that white participation rates are not explained by controls. For blacks, an upward shift (along with a change in gradient) results, so that their profile resembles that of whites. This suggests that racial equalisation occurs, once other characteristics are taken into consideration. It is problematic to emphasise this result, given that the constant adjustment was in this case very sensitive to the chosen cohort. Age effects (Figure 14.2) again remain stable for the youth, while a downward participation shift result for older groups. If one considers these together with the downward employment movements (Figure 13.2), it is evident that the changes in the latter dominate in explaining the increases in unemployment for older individuals.

Figure 12: Unemployment rate by cohort and their decompositions – with controls, 1995-2005

Figure 12.1 Unemployment rate by birth cohort and age

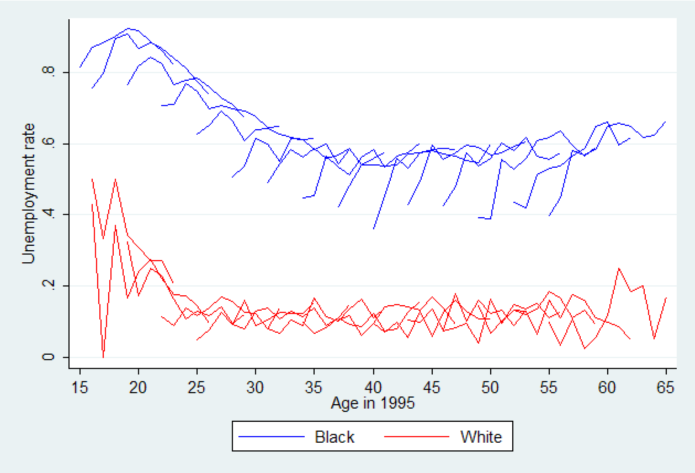


Figure 12.3 Unemployment rate birth cohort effects

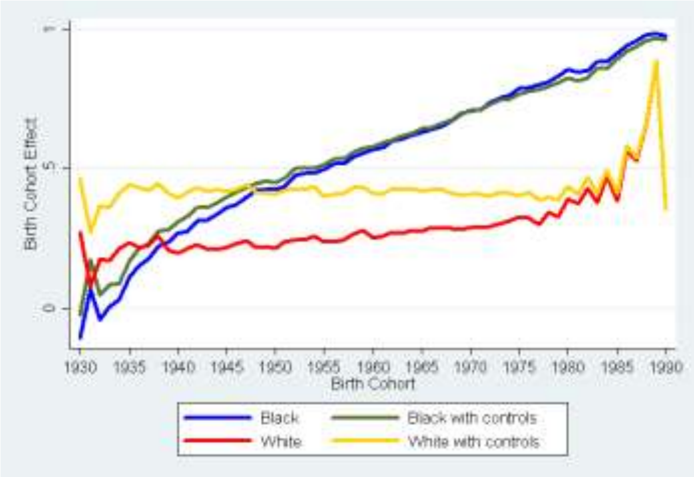


Figure 12.2 Unemployment rate age effects

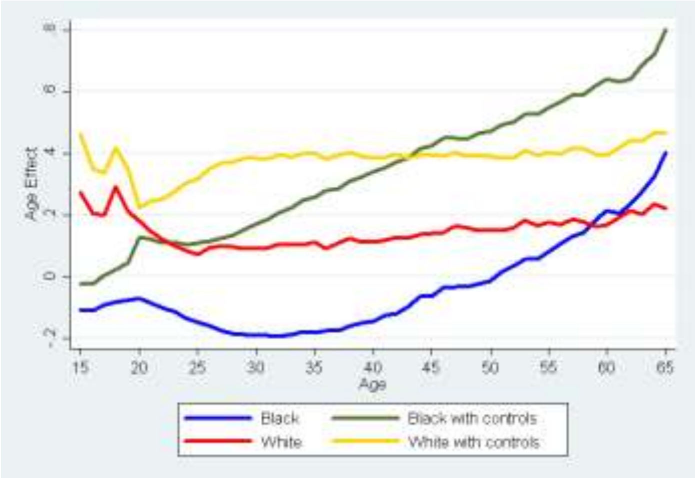


Figure 12.4 Unemployment year effects

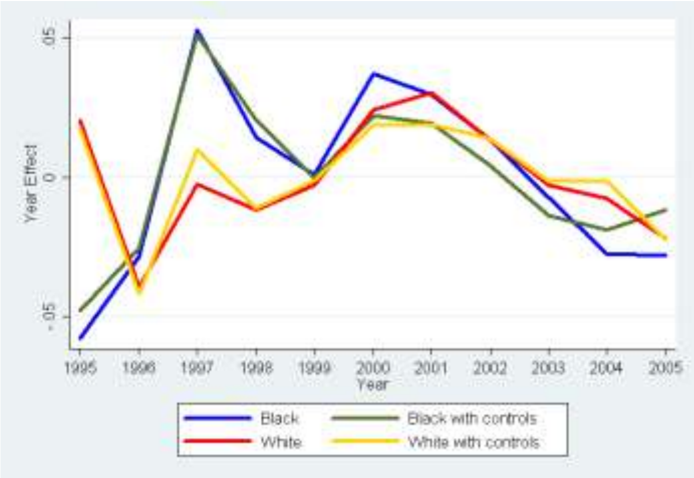


Figure 13: Employment rate by cohort and their decompositions – with controls, 1995-2005

Figure 13.1 Employment rate by birth cohort and age

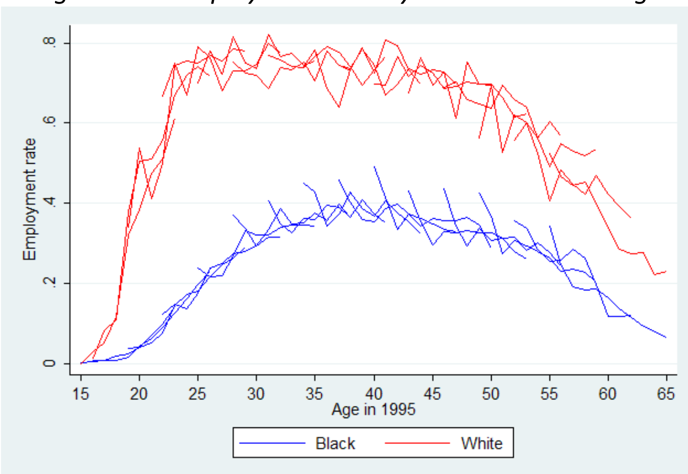


Figure 13.3 Employment rate birth cohort effects

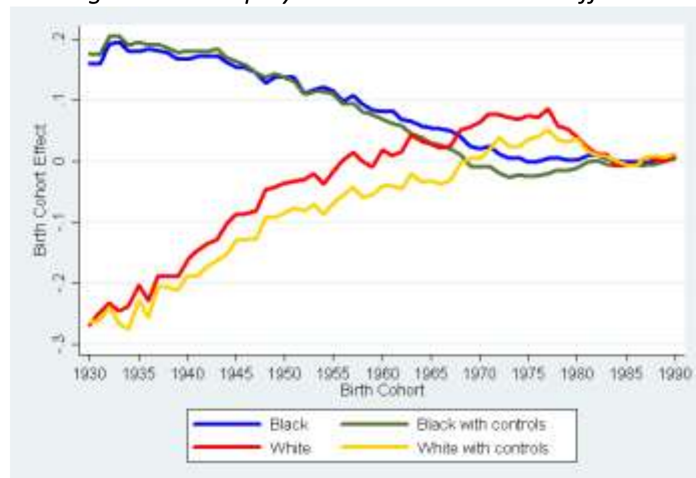


Figure 13.2 Employment rate age effects

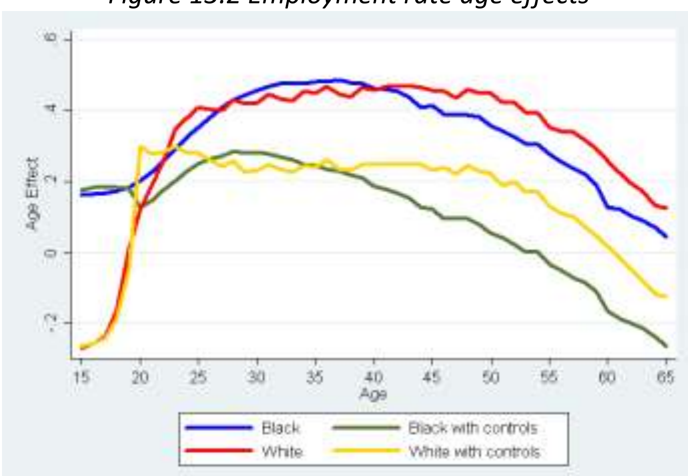


Figure 13.4 Employment rate year effects

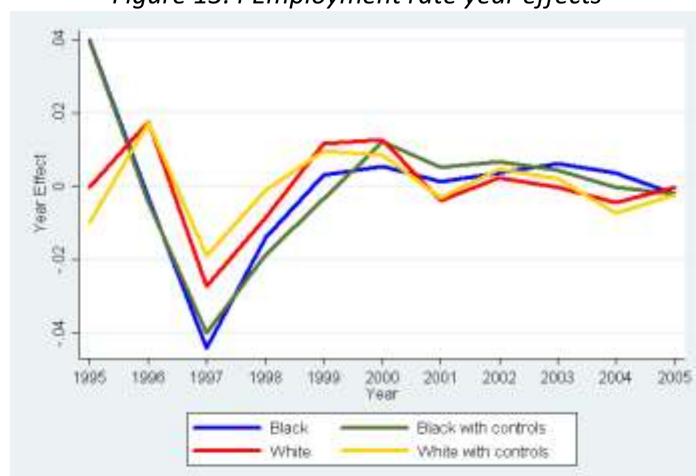


Figure 14: Participation rate by cohort and their decompositions – with controls, 1995-2005

Figure 14.1 Participation rate by birth cohort and age

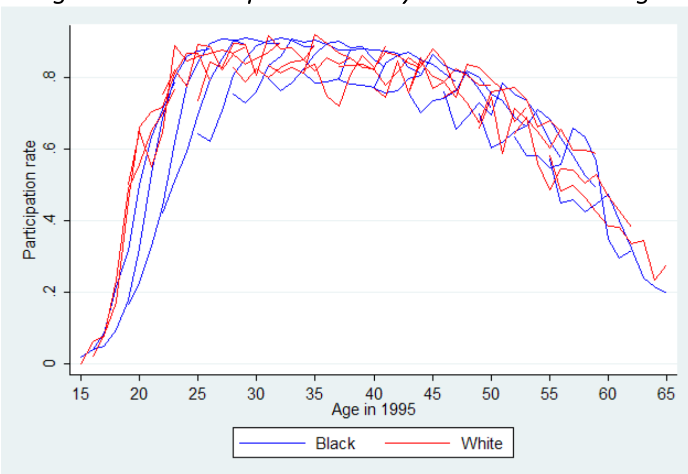


Figure 14.2 Participation rate age effects

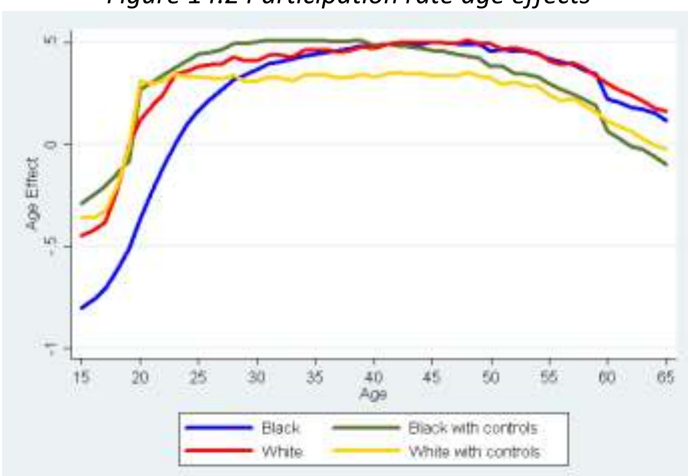


Figure 14.3 Participation rate birth cohort effects

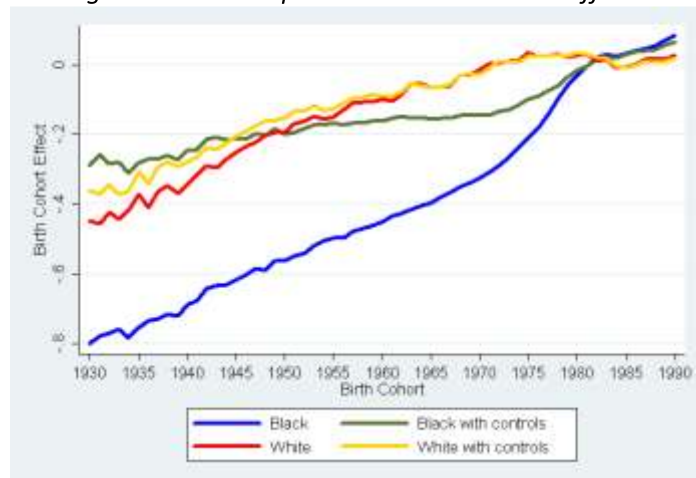
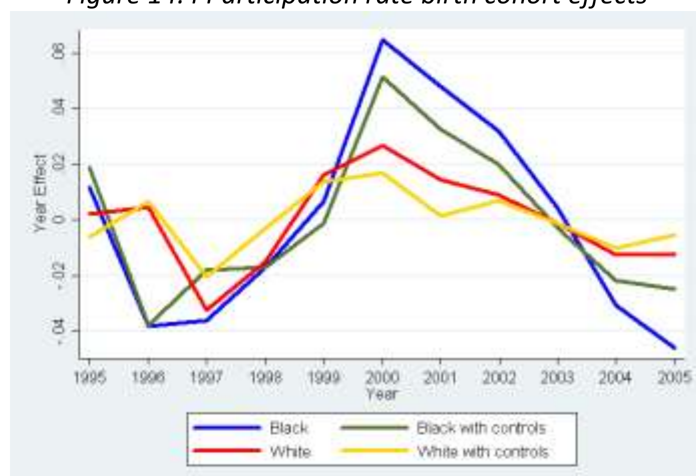


Figure 14.4 Participation rate birth cohort effects



4 Conclusion

This paper attempted to contribute to the South African unemployment debate by analysing the 17 available post-1994 household survey datasets at the cohort- rather than the individual level. This was done by means of a decomposition analysis that identified the between cohort and population group changes in the unemployment rate attributable to cyclical, generational and life-cycle effects. This analysis was also extended to a decomposition of the participation and employment rates. The decomposition indicates that the higher unemployment rates faced by the young are predominantly due to the disadvantage of entering the labour market more recently, rather than being attributable to their age. It was also shown that the bulk of this generational disadvantage was the result of an increase in the participation rate, rather than a decrease in employment opportunities. We find some correspondence between the cyclical variation in unemployment and the business cycle, although this relationship might be marred by survey-specific sampling error. Finally, we also estimated the decomposition regression while controlling for a set of observable characteristics. In most cases, this does not have a substantial effect on the shape of the age, birth cohort or year effect curves, which implies the presence of other important determinants, possibly at a macro-economic or regulatory level, that have an important influence on unemployment trends.

5 Bibliography

Baltagi, B. (2005). *Econometric Analysis of Panel Data 3rd Edition.* Chichester: John Wiley & Sons.

Branson, N. (2005). *The South African Labour Market 1995-2004: A Cohort Analysis.* Honours Thesis, University of Cape Town.

Casale, D., & Posel, D. (2002). "The Continued Feminisation of the Labour Force in South Africa: An Analysis of Recent Data and Trends. " *South African Journal of Economics* 70 (1):156-184.

Deaton, A. (1985). Panel Data from Time Series of Cross-Sections. *Journal of Econometrics, 30*, 109-126.

Deaton, A. (1997). *The Analysis of Household Surveys - A Microeconometric Approach to Development Policy.* Baltimore: Johns Hopkins University Press.

Devey, R., Skinner, C. and Valodia, I. (2006). "Second Best? Trends and Linkages in the Informal Economy in South Africa." *DPRU Working Paper* 06/102. Development Policy Research Unit, University of Cape Town.

Grün, C. (2004). *Racial and Gender Wage Differentials in South Africa: What can Cohort Data tell?* Available [Online]: http://www.wits.ac.za/sebs/downloads/cohort_data.pdf.

Hsiao, C. (2003). *Analaysis of Panel Data - Second Edition.* Cambridge: The Press Syndicate of the University of Cambridge.

Inoue, A. (2005). *Efficient Estimation and Inference in Linear Pseudo-Panel Data Models.* Texas A&M University, Department of Economics, Spring Workshop. Available [Online]: http://econweb.tamu.edu/workshops/Econometrics/Atsushi%20Inoue.pdf

Kingdon, G., & Knight, J. (2005). *Unemployment in South Africa, 1995-2003: causes problems and policies.* Retrieved from Global Poverty Research Group Working Paper 010. Available [Online]:http://www.gprg.org/pubs/workingpapers/pdfs/gprg-wps-010.pdf.

Kingdon, G., & Knight, J. (2006). The measurement of unemployment when unemployment is high. *Labour Economics* 13(3): 291-315

Mlatsheni, C. & Rospabé, S. (2002). Why is Youth Unemployment so High and Unequally spread in South Africa? *DPRU Working Paper* 02/65. Development Policy Research Unit, University of Cape Town.

Oosthuizen, M. & Bhorat, H. (2005). The post-apartheid South African labour market. *DPRU Working Paper* 05/093. Development Policy Research Unit, University of Cape Town.

Republic of South Africa, (1995). *Meeting the Commitment to Free and Compulsory General Education*. Chapter 13.White Paper on Education and Training. Notice 196 of 1995. Department of Education. Available [Online]: http://www.polity.org.za/html/govdocs/white_papers/educ5.html

South African Reserve Bank (SARB), (2006). *Quarterly Bulletin – June 2006 No240*. Pretoria: South African Reserve Bank.

Verbeek, M., & Nijman, T. (1992). Can Cohort Data be Treated as Genuine Panel Data? *Empirical Economics, 17*, 9-23.

Appendix: Figures

Figure A1: Unemployment rate by birth cohort and year

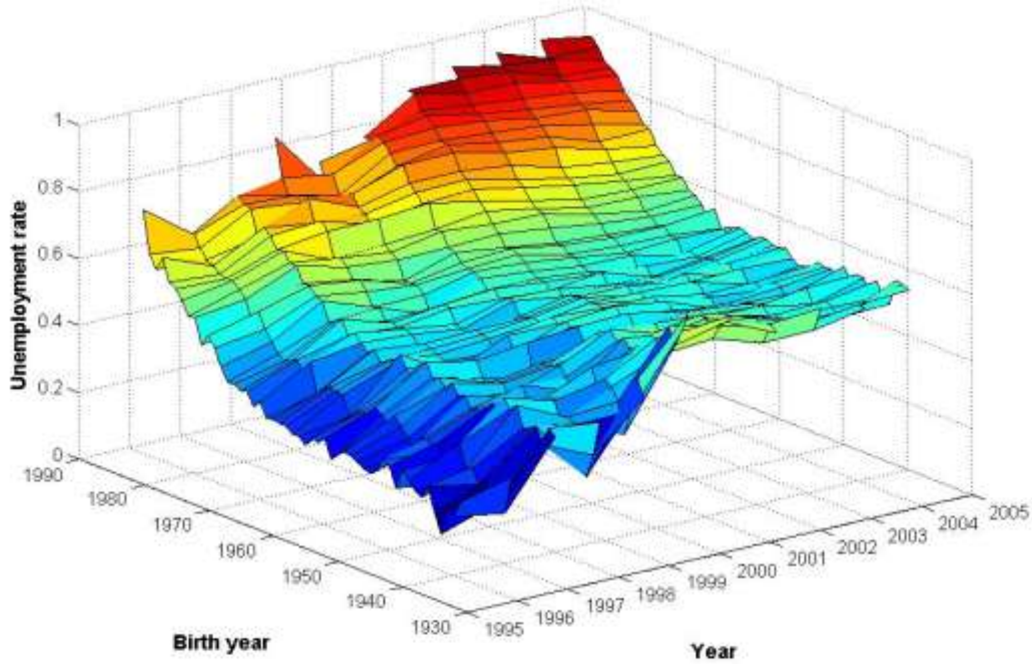


Figure A2: Employment rate by birth cohort and year

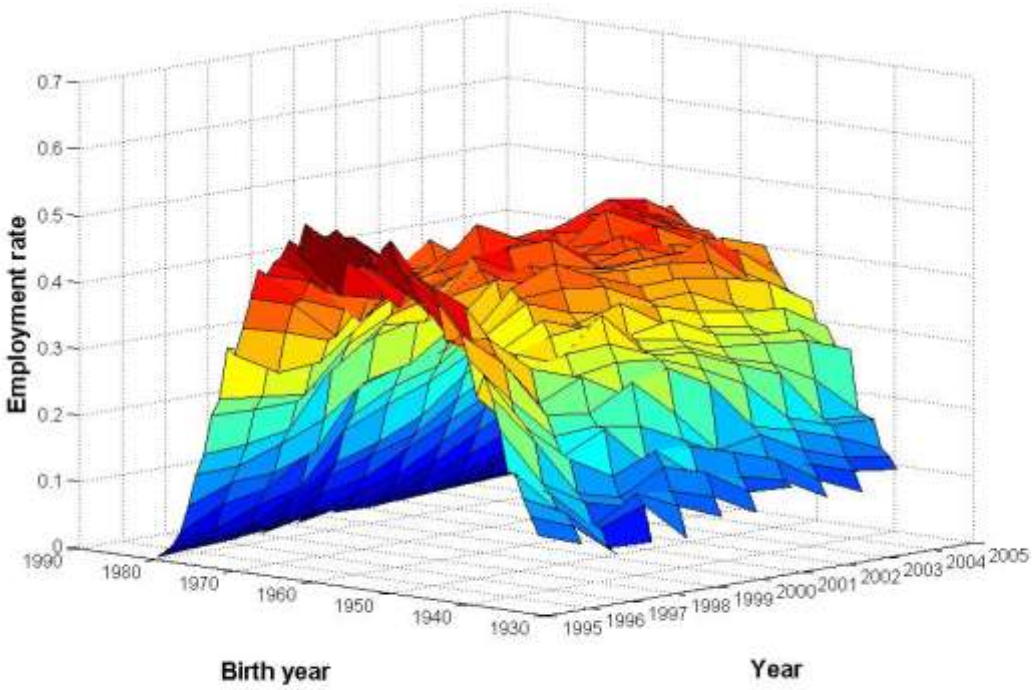


Figure A3: Participation rate by birth cohort and year